

Individual toddlers' interactions with teachers, peers, and the classroom environment in Danish and Dutch childcare: First validation of the inCLASS Toddler

Pauline L. Slot^{a,*}, Marinka M. Willemsen^{b,1}, Dorthe Bleses^c, Jason Downer^d

^a Department of Development and Education of Youth in Diverse Societies, Utrecht University, Utrecht, the Netherlands

^b TrygFonden's Centre for Child Research, Aarhus University, Aarhus, Denmark; Sardes, Utrecht, the Netherlands

^c TrygFonden's Centre for Child Research, Aarhus University, Aarhus, Denmark

^d Center for Advanced Study of Teaching and Learning, University of Virginia, Charlottesville, United States

ARTICLE INFO

Keywords:

Teacher-child interactions
Peer interactions
Task engagement
Behavior regulation
Toddler care

ABSTRACT

Toddlers' interactions with teachers, peers and the classroom environment are critical for their academic and social development. The newly developed measure – Individualized Classroom Assessment Scoring System for Toddlers (inCLASS Toddler) – evaluated the quality of toddlers' interactions in two European countries. This first validation study examined the structural, construct and criterion validity of the inCLASS Toddler. Using observational data of Danish and Dutch toddlers ($N = 211$) across 58 classrooms, the hypothesized four-domain structure showed an adequate and marginal fit to the Danish and Dutch data, respectively. Construct validity indicated within-country differences for children's age, but not for gender. Criterion validity was evidenced by small to large relations with children's social-emotional, language, and math skills. These results support the applicability of the inCLASS Toddler for observing toddlers' situated skills in interacting with their environment, which can be useful for practitioners and researchers to evaluate toddlers' interactions in an ecologically valid way.

1. Educational relevance

This study supported the applicability of an ecologically valid observational tool to assess toddlers' interactions with teacher and peers, as well as their levels of engagement and adaptive classroom behavior. Four different aspects of individual children's classroom experiences could be distinguished: teacher interactions, peer interactions, task orientation, and behavior control. The reported reliability and validity of the inCLASS Toddler support the usefulness of this observational tool for assessing individual children's skills for both research and practice, and highlights the added value compared to the use of overall classroom measures for interaction quality.

2. Introduction

Toddlerhood is characterized by rapid growth across several

domains of development and lays the foundation for future development and school success (e.g., National Scientific Council on the Developing Child, 2009). Children develop through positive, nurturing relationships with caregiving adults (e.g., Bronfenbrenner & Morris, 1998; Shonkoff, 2011) and these responsive interactions contribute to their cognitive, language, early math, and social-emotional development (Thompson, 2006). As the majority of toddlers is enrolled in center-based childcare, especially in West-European countries such as the Netherlands and Denmark (OECD, 2023), this environment with professional caregivers provides an important pedagogical context for growth and development depending on the quality of this environment (Melhuish et al., 2015). High quality childcare has shown to be positively related to children's cognitive, language, social-emotional development (e.g., Côté et al., 1998, 2000, 2006; Ruzek et al., 2014).

Commonly quality is conceptualized and measured at the classroom

* Corresponding author at: Utrecht University, Faculty of Social Sciences, Department of Development and Education of Youth in Diverse Societies, Heidelberglaan 1, 3584 CS Utrecht, the Netherlands.

E-mail address: p.l.slot@uu.nl (P.L. Slot).

¹ Co-first author.

<https://doi.org/10.1016/j.lindif.2023.102347>

Received 16 September 2022; Received in revised form 3 August 2023; Accepted 9 August 2023

Available online 29 August 2023

1041-6080/© 2023 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

level. Although informative, this does not adequately capture individual children's experiences (Burchinal, 2018; Downer et al., 2010). Increasing evidence suggests that overall classroom quality might affect children differently, for instance depending on child characteristics (Rimm-Kaufman et al., 2009; Vitiello, Booren, et al., 2012). For example, two studies involving Dutch two- and three-year-old children revealed that high observed emotional and behavioral support in the classroom positively affected children's social skills, but only for children with low behavioral self-regulation skills (Broekhuizen et al., 2016, 2015). Moreover, existing classroom quality measures, such as the Classroom Assessment Scoring System, (CLASS Toddler; La Paro et al., 2012) or Early Childhood Environment Rating System (ECERS; Harms et al., 2015), only provide an overall and generalized measure of quality which may mask individual differences in children's classroom experiences and their interactions within the broader classroom context. In addition, these measures emphasize the teachers' role (Connor et al., 2020), while neglecting the role of peers as an additional source of socialization considered important to capture when investigating classroom quality (Fabes et al., 2003). Peer relations have shown to contribute to children's development, such as social skills (e.g., Williams et al., 2010) and language skills (e.g., Chen et al., 2020). Higher quality of interactions of both teachers and peers, from an individual child's perspective, were related to higher academic achievement, more adaptive learning behavior and better self-regulation (Fantuzzo et al., 2005; Gifford-Smith & Brownell, 2003; Hamre et al., 2007; Pianta & Stuhlman, 2004). Also, children's initiative and engagement in exploring their environment plays a role in the extent to which these interactions with teachers and peers are conducive to development and learning as children are viewed as active participants in their own development (Thompson, 2006). For example, children with better task-related behavior showed better regulation of emotions and gains in academic skills (Fantuzzo et al., 2004; McClelland et al., 2000).

Given the importance of children's interactions with teachers, peers, and their classroom environment, it is essential to capture the quality of these interactions from the perspective of individual children. For preschoolers, such an observational tool exists, the Individualized Classroom Assessment Scoring System Pre-Kindergarten (inCLASS Pre-K; Downer et al., 2010), which was originally developed in the United States (U.S.), but validated in several other countries, including Denmark (Slot & Bleses, 2018). The current study introduces a newly developed measure to capture *toddlers'* individual experiences in the classroom, the inCLASS Toddler, and presents a first validation of this measure within Denmark and the Netherlands.

2.1. Quality of Individual Interactions with the InCLASS Pre-K

The inCLASS Pre-K taps into four different domains of interactions: Teacher Interactions, Peer Interactions, Task Orientation, and Conflict Interactions (Downer et al., 2010). The first domain, Teacher Interactions, consists of two dimensions: Positive Engagement evaluates the child's emotionally close and secure relationship with the teacher; Teacher Communication captures the degree to which the child initiates and maintains verbal interactions with the teacher. The second domain, Peer Interactions, comprises three dimensions: Peer Sociability evaluates the child's interest in peers, and positive emotions and behaviors directed towards peers; Peer Communication reflects the degree to which the child initiates and maintains verbal interactions with peers; Peer Assertiveness captures the extent to which the child initiates contact or play with peers and demonstrates confidence and leadership in this contact. The third domain is Task Orientation, which includes two dimensions: Engagement with Tasks reflects the extent to which the child shows consistent and active engagement in tasks and activities; Self-Reliance captures the degree to which the child takes learning into their own hands. The fourth domain, Conflict Interactions, consists of three dimensions: Teacher Conflict and Peer Conflict reflect the degree to which the child shows negativity or aggression towards the teacher or

peers; Behavior Control (reversely coded) measures the degree to which the child regulates behavior to match the classroom expectations.

U.S. studies using the inCLASS Pre-K have demonstrated associations with child outcomes. For example, Sabol et al. (2018) showed the added value of the quality of individual children's interactions above and beyond the overall classroom quality. Higher quality interactions with teachers predicted literacy skills, and the quality of peer interactions was related to better language and self-regulation skills. Further, better task orientation was related to closer teacher relationships as measured with the Student-Teacher Relationship Scale (STRS; Kim et al., 2019; Pianta et al., 2001). In another study, children's interactions with teachers were related to their compliance and executive function skills, whereas children's task orientation was associated with their emotion regulation (Williford et al., 2013). Children's task engagement was also found to be related to language skills (Bohlmann & Downer, 2016) and self-regulation skills based on teacher ratings (Kim et al., 2019). Children's interactions with the teacher and with peers also showed to be related to their social skills based on teacher ratings (Kim et al., 2019). Finally, another study demonstrated how different inCLASS profiles mediated the association between the child's relation with the teacher and their expressive vocabulary skills (Rojas & Abenavoli, 2021).

To conclude, the inCLASS Pre-K has four underlying domains, which have revealed differentiated patterns of associations with child development covering a broad set of skills considered important for school readiness. The inCLASS Pre-K version reflects the developmental phase of preschoolers with a strong verbal language orientation and basic self-regulation abilities. However, toddlerhood reflects a different developmental phase requiring adaptations to the inCLASS Pre-K to make it suitable for toddlers.

2.2. Adaptation of inCLASS Pre-K to the InCLASS Toddler

The inCLASS Pre-K taps into fundamental aspects of individual children's classroom experiences, but the manifestation of these experiences may be different for toddlers. Therefore, we maintained the overarching domain structure of the inCLASS of Teacher Interactions, Peer Interactions, Task Orientation and Conflict Interactions (yet framed the latter domain positively and labeled it Behavior Control) but adapted some of the behavioral markers in the dimensions to better align with toddlers' developmental stage. Also, we modified the examples in the Appendix of the inCLASS manual to reflect more age-appropriate classroom experiences for toddlers.

During toddlerhood children take the first steps in gaining autonomy within connected relations with others, as they balance their need for independence with the need for security and comfort from adults (Bronson, 2000; Calkins, 2007; Sroufe, 1996). Guedes et al. (2020) showed that individual children indeed had more positive interactions with teachers in classrooms where teachers were responsive and child-centered thus reflecting a high emotional climate. Supportive interactions are based on a continuous give-and-take or "serve and return" exchanges, which begin to develop in infancy (Thompson, 1999). However, there is large individual variation in early language development (e.g., Bleses et al., 2008; Hoff, 2013), which means that children's interactions also (partly) include non-verbal (e.g., tracking, making eye contact) and pre-linguistic behaviors (e.g., gesturing, making sounds) as they engage with teachers and peers. Therefore, non-verbal, and pre-linguistic behaviors were added to the inCLASS indicators specifically for the dimensions Positive Engagement with the Teacher, Teacher Communication, Peer Sociability, Peer Communication and Peer Assertiveness.

Furthermore, as attention span is still developing at this young age (e.g., Garon et al., 2008), we adapted the dimension Engagement within Tasks and added whether a child is able to continue playing or regain focus in routines and activities after a short distraction. Although, self-regulation skills develop rapidly, displaying compliance, as a precursor of self-regulation, is considered more age-appropriate in toddlerhood

(Kopp, 1982). Therefore, we added 'Compliance' as an indicator within the dimension Behavior Control, referring to the child's ability to follow the rules and structure of the classroom. A complete overview of adaptations made to the inCLASS Pre-K is provided in Appendix A.

Next, the adapted measure was used in a pilot with videos of Danish toddlers across different activities, such as routines and play. The first and third author extensively discussed the findings of this pilot and made final adaptations to the measure that were reviewed by the fourth author, which resulted in the current version of the inCLASS Toddler that was used in Denmark and the Netherlands. Although the teaching practices and structure of classrooms may differ cross-culturally, thus offering different kinds of opportunities for children, the types of children's individual interactions (i.e., with teacher, peers, and the classroom environment) are expected to remain similar. As such, an observational tool like the inCLASS Toddler should be applicable across cultural contexts. The current study provides a first validation of this assessment tool in Denmark and the Netherlands.

2.3. Framing the psychometric properties of the inCLASS Toddler through the Pre-K

Fundamental classroom interactions with teachers, peers, and tasks show continuity throughout childhood, yet their specific behavioral markers may differ (Hamre et al., 2007). Therefore, the adaptation of an observation instrument like the inCLASS to toddlerhood requires the testing of the psychometric properties, including structural validity (dimensionality), construct validity, and criterion-related validity.

The inCLASS Pre-K showed a four-factor structure in several studies in Austria, Denmark, Germany, and the U.S., albeit with lower factor loadings for the conflict dimensions (Booren et al., 2012; Downer et al., 2010; Schmidt & Embacher, 2021; Slot & Bleses, 2018; Vitiello, Moas, et al., 2012; Von Suchodoletz et al., 2015; Williford et al., 2013). Despite the restricted variance of the conflict dimensions, this variability can still be meaningful and relevant. For instance, boys showed more conflictual interactions than girls in Denmark, Germany, and the U.S. (Downer et al., 2010; Slot & Bleses, 2018; Von Suchodoletz et al., 2015) and in Danish preschools also older children showed more conflicts as compared to younger children (Slot & Bleses, 2018). Thus, prior research supports the inclusion of the conflict dimensions despite the lower variability and therefore these were kept but scaled positively and included in the domain now called Behavior Control instead of Conflict Interactions.

Several studies on the construct validity of the inCLASS Pre-K support the measure's sensitivity to children's age and gender. Overall, older children score higher in terms of peer interactions and classroom engagement (Downer et al., 2010; Vitiello, Moas, et al., 2012). In a recent study by Slot and Bleses (2018), contrasting evidence was found at first, such that older preschool children scored lower on peer interactions and classroom engagement than their younger peers. Yet upon closer analysis, this age effect mostly disappeared when accounting for the classroom age composition. The age heterogeneity was substantial in Danish preschools and especially for older children, the lack of same-aged or older peers may have resulted in fewer positive peer interactions, more conflicts and less overall engagement. Concerning gender, girls score higher on peer interactions in preschool than boys (Von Suchodoletz et al., 2015) and girls have fewer conflictual interactions than boys (Downer et al., 2018; Slot & Bleses, 2018; Von Suchodoletz et al., 2015). However, in another study no gender differences were found regarding the quality of individual children's interactions (Downer et al., 2010).

Regarding criterion validity, positive associations were found between the quality of children's interactions and their language, literacy, and self-regulation skills (Bohlmann & Downer, 2016; Downer et al., 2010; Sabol et al., 2018; Williford et al., 2013). Moreover, examining three different profiles of overall positive engagement, typical engagement and negative engagement based on the inCLASS, Williford et al.

(2013) demonstrated differential relations with child outcomes. Positively engaged children performed better on self-regulation skills, whereas negatively engaged children made less gains in self-regulation and literacy.

To summarize, there is support for the psychometric rigor of the inCLASS Pre-K, including a four-factor structure confirmed across several countries and contexts, and evidence for construct and criterion validity. The current study adds to the existing knowledge by evaluating the psychometric properties of the newly developed inCLASS Toddler in a Danish and Dutch childcare context.

2.4. Early childcare in Denmark and the Netherlands

The childcare context in Denmark and the Netherlands is quite comparable in some respects but varies regarding other aspects (Bleses et al., 2020; Slot, 2018). The enrollment rate is about 80 % in the Netherlands, though part-time for on average two days per week, and about 90 % in Denmark. The teacher-to-children ratio is 1:3.5 in Denmark, whereas in the Netherlands the ratio is dependent on the age composition of the classroom, with 1:5 for 1–2 years, 1:6 for 2–3 years, and 1:8 for 3–4 years. Teachers in Denmark are on average higher educated as about 60 % have a bachelor's degree compared to about 21 % in the Netherlands. Both countries share a holistic child-centered and strongly play-based approach, albeit with a slightly stronger emphasis on school readiness in the Netherlands (Bleses et al., 2020; Slot, 2018). Pedagogical plans are formulated on the level of the childcare center based on a national learning curriculum in Denmark and the Child Care Act (2005) in the Netherlands. Given the similarities in quality and pedagogy, it is interesting to investigate the experiences of individual children and evaluate the applicability of the measure in these two contexts.

2.5. Current study

Altogether, evidence from preschoolers showed the added value of the perspective on individual children's interactions as it contributes to their development in various domains. The observational tool inCLASS Pre-K has proven useful in capturing children's classroom experiences. In the present study we introduce an adaptation of the inCLASS Pre-K to measure individual children's interactions with teachers, peers, and the classroom environment in toddlerhood in Denmark and the Netherlands. We investigate three research questions: 1) What is the factor structure of the inCLASS Toddler (i.e., structural validity)? 2) Are there differences in individual children's interactions based on gender and age (i.e., construct validity)? 3) What are the associations between individual children's interactions and their language, math, and social-emotional skills (i.e., criterion validity)? The latter research question was only investigated for Danish children as there were no other child data available for the Dutch children.

Based on the evidence with the inCLASS Pre-K, we expect to find a similar four-factor structure with the four domains: Teacher interactions, Peer interactions, Task orientation and Behavior Control (in some papers referred to as Conflict interactions) (e.g., Booren et al., 2012; Downer et al., 2010; Slot & Bleses, 2018; Vitiello, Moas, et al., 2012; Von Suchodoletz et al., 2015). Concerning the second research question, we expected older children to score higher on peer interactions and task orientation (Downer et al., 2010; Vitiello, Moas, et al., 2012) and we expected girls to score higher on peer interactions as well as behavior control (Downer et al., 2010; Slot & Bleses, 2018; Von Suchodoletz et al., 2015). Finally, regarding the third research question, we expect to find positive associations between Danish toddlers' individual interactions and their language, math, and social-emotional skills, alike results with the inCLASS Pre-K (e.g., Bohlmann & Downer, 2016; Downer et al., 2010; Kim et al., 2019; Sabol et al., 2018; Williford et al., 2013).

3. Method

3.1. Participants

Data from two European countries were used in this study for a cross-country analysis. For an overview of descriptive characteristics of the Danish and Dutch samples, see Table 1.

3.2. Danish sample

The current study used a subsample of a large, randomized control trial “Play and Learn”, which aimed to strengthen children’s language and math skills by elevating the instructional and process quality in toddler classrooms (Bleses et al., 2020). The collection and treatment of all data for the project was approved by the Danish Data Protection Agency (approval no. j.nr. 2014-54-0822). All these daycare centers were located in a large southern city of Denmark. Teachers at the daycare centers had on average 11 years of experience and around 90 % held a BA-level degree in teaching as this was an inclusion criterium for the randomized control trial (Bleses et al., 2020). All teachers and parents were informed about the study with materials provided in five different languages and were assured that all data would be treated anonymously and confidentially. The subsample consisted of 99 children within 30 classrooms of 22 daycare centers. The selected centers were approached by the third author based on existing contact. At the time of the inCLASS Toddler observations, the Danish children were on average 22.75 months old ($SD = 4.68$, 13–33). Normality tests did not show significant skewness or kurtosis values for age (Skewness = 0.267, Kurtosis = 0.180, $p = 0.209$). The average staff-to-children ratio was 1:3.5, based on the live coding of the inCLASS in the classrooms, which is in line with the national regulations. The average group size was small with a mean of 6.54 children (see descriptive information in Table 1).

3.3. Dutch sample

The Dutch data were collected as an extension of the EU-funded CARE study into quality and effectiveness of Early Childhood Education and Care in Europe (Cadima et al., 2023; van de Riet & Slot, 2017). The data collection and treatment of all data was approved by the

Table 1
Descriptive characteristics of Danish and Dutch samples during inCLASS Toddler observations.

Variables	Denmark	The Netherlands
Children’s age	22.75 months	35.85 months
Children’s gender (girls-boys)	45 %–55 %	54 %–46 %
Mean number of teachers during observation cycle ^a	1.86	1.46
Mean number of children during observation cycle ^a	6.54	7.82
Teacher part of the activity ^a	67 %	95 %
Activities ^b		
Mealtime	25.64 %	28.80 %
Free play	29.23 %	23.20 %
Educational	5.64 %	27.20 %
Creative	5.13 %	20.80 %
Outdoor activities/play	28.21 %	0 %
Other	6.15% ^c	0 %

Mealtime: eating and drinking, Free play: indoor and outdoor free play, Educational: organized activities, such as shared book reading, Creative: organized activities, such as arts or crafts, Outdoor: play or activities outside or Other: routines or transitions. Number of teachers and children, teacher part of the activity and activity setting was collected by inCLASS Toddler coders during observations.

^a On child level.

^b On observation cycle level.

European Commission following the European Commission’s ethical regulations effective at that time (project number FP7-SSH-2013-2). The Dutch centers had to meet several inclusion criteria: centers had to meet national guidelines concerning structural characteristics (e.g., ratio, teachers’ pre-service qualifications) and classrooms had to serve mainly, though not exclusively, two-year-olds (Wyslowska & Slot, 2020). Centers were selected to cover regional variation and varying degrees of urbanization: there were centers from two (out of four) of the largest cities, centers from middle- to large-sized cities (as part of the national network), and centers from smaller towns across the country. Two types of centers were included: full-day childcare programs serving children from working parents (mainly reflecting Dutch middle-class families) and half-day programs aimed at disadvantaged children (e.g., with lower educated parents and/or from non-Dutch backgrounds). These two types of programs represent relevant variation in the Dutch context (van de Riet & Slot, 2017). All centers were recruited by phone by the first author. All parents and teachers were fully informed about the purpose of the study and were assured that all data would be treated anonymously and confidentially. They were only included in the study after obtaining active written consent. 26 % of the teachers had a bachelor’s degree and the remainder had an intermediate vocational degree (in line with Dutch quality regulations). Teachers had on average 15 years of work experience. The sample of the inCLASS Toddler observations consisted of 114 Dutch children within 28 groups of 10 daycare centers. The Dutch children were on average 35.85 months old ($SD = 4.44$, 21–42). Normality tests showed significant skewness and kurtosis values for age (Skewness = 0.008, Kurtosis = 0.967, $p = 0.039$), indicating the age distribution is dominated by older children. The mean group size 15.41 ($SD = 2.96$) and the teacher-to-children ratio is 1:7.42.

3.4. Procedures

3.4.1. Data collection

For the Danish sample, all measures were collected in 2015 at the end of summer until the beginning of autumn. The inCLASS Toddler scores were collected at the beginning of that period through a live coding procedure carried out by three Master students. They visited classrooms with the aim of collecting four observations from on average four randomly selected children from the class list through consecutive observation cycles alternating between children. If children were absent from a particular classroom on the day of the observation, more children were selected from another classroom to compensate, resulting occasionally in more or less than four children selected from a single classroom. During the observation, observers also noted the activity and how many children and adults were present. The noted down Danish activities were later categorized into one of 5 categories commonly distinguished in toddler classrooms (Slot et al., 2015; Guedes et al., 2020): Mealtime (e.g., fruit moment, lunch), Free play (indoor or outdoor), Educational (e.g., shared book reading, motor skills activity), Creative (e.g., clay modelling, singing), Outdoor play or activities or Other (e.g., routine activities such as washing hands, queueing to go outside). To limit the impact of observations on the teachers, no specific activities were requested for observation. Most observations concerned free play (indoor and outdoor) and meals, and to a lesser degree organized educational or creative activities. This distribution of activities reflects the common practice in Danish daycare and was also captured by other observational data of the “Play and Learn” study. Only for the Danish sample, pre-test intervention scores for social-emotional, language, and math skills were included, which were collected 1.3 months after the inCLASS Toddler data collection.

For the Dutch sample, four Master students made video recordings on two morning visits in classrooms following the procedures of the extension of the EU-study CARE in 2015 from the beginning of February until the beginning of July. Teachers were asked about the program of the day in advance, and on the day itself again, in order to adjust the timing of the video recording to the planned activities. Teachers were

not given any instructions and were requested to carry out their daily schedule as usual. Four video recordings were made in each classroom covering the daily activities as they occurred naturally during the two visits (Cadima et al., 2023; Guedes et al., 2020): Free play (indoor only), Mealtime (fruit or snack), Educational (e.g., shared book reading, making puzzles), and Creative activities (e.g., arts and crafts). Coding with the inCLASS Toddler occurred afterwards for all children that were present in at least two out of four videotaped activities based on a first screening of the videos by an assistant. On average four children were coded per classroom. Occasionally if less than four children were present in multiple videos per classroom, this was compensated by additional children from another classroom. Coders also noted the type of activity and how many children and adults were present.

For both the Danish and Dutch samples, children were only coded with the inCLASS Toddler if they were present during at least 8 min of the observation. Observation with the inCLASS Toddler took place in 15-min cycles: 10 min of observation and 5 min of coding while alternating between the selected children. As recommended by the manual, multiple cycles were collected for every child (DK: M = 2.9; NL: M = 3.2).

3.4.2. Training, coding, and inter-rater reliability

The inCLASS Toddler was used to assess individual children's classroom interactions in the Danish and Dutch daycare settings. An official two-day training was provided by the first author who is one of the developers of the inCLASS Toddler. The trainer introduced the dimensions, the indicators and behavior markers of the inCLASS Toddler. After each dimension was extensively discussed and illustrated with a short videoclip, the observers practiced the coding procedures of at least 5 videos of children in different activity settings (e.g., free play, mealtime, and organized group activities). Thereafter, observers were required to reach a minimal reliable coding score of 80 % on a reliability test of five master-coded videos. Coding was considered reliable when a code was within one point with the first author's coding.

To establish inter-rater reliability, Danish and Dutch observations were independently double-coded, which was respectively 22.2 % (n = 22) and 33.3 % (n = 38) of the child observations. For the Danish observations, double-coding took place in a live situation with two coders present at the observation, while for the Dutch observations double-coding took place via video recordings. Teams of different coders were used to double-code the observations, therefore a one-way random-effects model with absolute agreement for single measures was conducted in STATA Version 16 using the intraclass correlations (ICC) command. For the Danish sample, Overall, results show that the ICCs for the Danish scores ranged from good (>0.75 - <0.9) to excellent (>0.9) (Koo & Li, 2016). The inter-rater reliability of the Dutch scores ranged from poor (<0.5) to excellent (>0.9), see Table 2 for the ICCs per dimension. Besides ICCs, which indicate correlations of ratings made on the same targets, percentage of coding agreement within one scale point is also given in Table 2. Some dimensions showed limited variation, which may explain the lower ICCs in case differences were found for the Dutch toddlers on the conflict dimensions (Teacher conflict, Peer conflict, and

Behavior Control) as the percentage agreement within one scale point was substantially better.

3.5. Measures

3.5.1. inCLASS Toddler

Children's interactions in the inCLASS Toddler are measured on 10 dimensions: Positive Engagement with the Teacher, Teacher Communication, Teacher Conflict, Peer Sociability, Peer Communication, Peer Assertiveness, Peer Conflict, Engagement with Tasks, Self-Reliance and Behavior Control (Slot et al., 2016). *Positive Engagement with the Teacher* reflects the degree to which the child is emotionally connected to the teacher(s) and uses the teacher as a secure base. *Teacher Communication* measures the extent to which the child initiates and maintains conversation with the teacher using (non)verbal and prelinguistic communication. *Teacher Conflict* refers to tension, resistance, or negativity in the child's relation with the teacher(s). *Peer Sociability* assesses the extent to which a child shares positive emotions and behaviors with peers. *Peer Communication* reflects the degree to which the child initiates, joins or maintains conversation with peers using (non)verbal and prelinguistic communication. *Peer Assertiveness* measures the degree to which the child uses takes initiative and shows leadership in contact with peers. *Peer Conflict* refers to tension, resistance, or negativity in the child's relation with peers. *Engagement with Tasks* measures the degree to which the child is involved in play and activities, including the level of intensity, concentration or enthusiasm displayed, the proportion of time the child spends on play or activities and whether a child can continue after a distraction. *Self-Reliance* reflects the degree to which a child takes learning into their own hands, for instance through exploration and novelty seeking and showing confidence and persistence in doing so. *Behavior Control* measures the extent to which the child is able to meet the classroom expectations and show compliance.

Each dimension is rated on a 7-point scale with 1 or 2 meaning that the child scores low on that aspect, 3, 4 or 5 meaning that the child scored in the midrange, and 6 and 7 meaning that the child scored high on that aspect. The Appendix of the inCLASS Toddler manual provides elaborate examples of children's behavior across different activity settings to guide observers' coding. Table 3 provides some examples, taken from the Appendix of the manual, to illustrate the differences between a score in the midrange and a high score.

3.5.2. Social-emotional skills

The Social-Emotional Assessment/Evaluation Measure (SEAM) was used to measure toddlers' social-emotional skills and competencies across ten domains (Squires et al., 2014). For this study, a Danish adaption of SEAM is used of which the ten domains are divided into the Empathy index and the Self-Regulation & Cooperation index (Sjoe et al., 2017). The Empathy index includes six domains: healthy interactions, expression of emotion, regulations of socio-emotional response, empathy, sharing and engaging, and self-image (test-true score correlation; girls = 0.86–0.87; boys = 0.82–0.91). An example item is: "Toddler tries to comfort others when they are upset". The Self-Regulation & Cooperation index consists of four domains: independence, regulation of attention and activity level, cooperation, and adaptive skills (test-true score correlation; girls = 0.80–0.86; boys = 0.82–0.84). An example item is: "Toddler cooperates with simple requests". In total, 35 items are rated by a teacher on a scale from 'Very true' (3), 'Somewhat true' (2), 'Rarely true' (1), and 'Not true' (0).

3.5.3. Language skills

Language skills of the Danish toddlers in this study were assessed by means of a Danish adaptation (Vach et al., 2010) of the MacArthur-Bates Communicative Development Inventories (CDI; Fenson et al., 2007), called the CDI-Educator (Bleses et al., 2018). The CDI-Educator is a teacher-reported scale which measures children's early vocabulary skills by means of two language reports on (1) the child's vocabulary and (2)

Table 2
Inter-rater reliability indicated by ICC values and percentage of agreement.

Dimensions	Reliability (ICC)		Agreement (%)	
	DK	NL	DK	NL
Pos. Engagement w. Teacher	0.94	0.77	100	82
Teacher Communication	0.97	0.81	100	71
Teacher Conflict	0.98	0.42	100	97
Peer Sociability	0.95	0.68	100	84
Peer Communication	0.95	0.79	100	79
Peer Assertiveness	0.96	0.69	100	66
Peer Conflict	0.95	0.26	100	97
Task Engagement	0.91	0.70	97	82
Self-Reliance	0.93	0.73	97	74
Behavior Control	0.88	0.39	100	71

Table 3
Examples of inCLASS Toddler dimensions in the mid- and high range from the inCLASS Toddler Appendix.

Dimensions	Mid-range	High range
Positive engagement with teacher	1) The child responds to the teacher's requests ("everyone stands up! Everyone pretend to be a bear!") and shares smiles and laughter with peers but does not look at the teacher. 2) The child runs to sit next to the teacher at the beginning of story time but then scoots over near to a peer a few minutes later.	1) The child stands next to the teacher and puts her arms around her while she is reading a story to the class. 2) The child smiles every time (s) he and the teacher make eye contact. 3) The child smiles and laughs when the teacher is smiling and laughing.
Peer communication	1) The child tells peers at the art table to look at his painting saying "Look" several times but does not initiate or sustain any further interaction. 2) The child starts playing next to some children gesturing or saying a few words to his peers about this game.	1) The child describes his drawing to peers and shows interest in other peers' drawings either by pointing or by using language. 2) When two peers are playing next to each other the child starts to play the same game while communicating verbally or non-verbally with the two other peers.
Self-reliance	1) When playing with clay the teacher has to demonstrate how to make a ball or flatten the clay, but after that the child can do it himself.	1) In painting the child discovers by himself that he can mix paint to create new colors. 2) While playing with clay, all children make balls, and this child discovers that he can make snakes out of the balls by rolling them gently.

the child's use of words. The first report deals with 70 vocabulary items in 9 topics, such as words about animals and things, people and routines or body parts. Vocabulary scores were calculated by adding up the words produced by the child. The second report assesses in 5 questions the extent to which a child uses language to refer to objects, actions or people distant from situation at hand, such as talking about experiences, people or objects that are not present. Language use was scored on a 4-point scale with the following categories (1) "not yet", (2) "sometimes", (3) "often", or (4) "always". Test-retest correlations are 0.68 for vocabulary and 0.54 for language use and internal consistency measures demonstrate reliability (Cronbach's alpha: 0.98 for vocabulary and 0.88 for language use, see blinded for review, for the validation study).

3.5.4. Math language and numeracy skills

Math language and numeracy skills of the Danish toddlers in this study were measured by a research-developed checklist that contains two subtests on (1) numeracy development and (2) content-specific math language that supports math development (Bleses et al., 2020). The checklist is teacher-reported and consists of 41 items. The first subtest includes 13 items that measure toddlers' early numeracy skills (numbers and counting skills) and the second subset includes 28 items which measure children's comprehension and use of math language (words for sizes, quantities, shapes, and space). Scores reflect the presence of a specific skill by noting down if the math skill is produced (1) "not yet", (2) "sometimes", (3) "often", or (4) "always". Psychometric tests showed high internal validity with Cronbach's alpha values around 0.95 and item-total correlations above 0.50, and substantial criterion validity with the CDI-measures with correlations between 0.60 and 73 (Bleses et al., 2020).

3.6. Analysis strategy

Before analysis, observation cycle scores were aggregated to the child level. Toddlers with only 1 cycle of data collection were excluded

before this process, resulting in 97 observations for the Danish data and 114 for the Dutch data. Assumptions of linearity (i.e., linear relationships between variables) and normality were tested. As expected, mostly scores for the conflict dimensions violated linearity and were skewed (see also, Downer et al., 2010).

To address the first research question of investigating the hypothesized four-factor structure, a Confirmatory Factor Analysis (CFA) was conducted in Mplus Version 8, separately for the Danish and Dutch context. To account for non-normality in the conflict dimensions, MLR was used as estimator in the CFA models. Model fit was evaluated on the basis of the comparative fit index (CFI), Tucker-Lewis index (TLI) and root mean square error of approximation (RMSEA). The CFI and TLI were considered adequate to good when the values were 0.95 and up and the RMSEA when the value was 0.06 or lower (Geiser, 2013). To improve model fit, modifications were made by examining modification indices from the model output. From there, a modification was chosen if it had (a) the highest potential to decrease the chi-square value, (b) at least a chi-square decrease of 3.84 (which corresponds with 1 degree of freedom at an alpha level of 0.05) and (c) was of theoretical relevance. If all conditions applied, the modification was incorporated into the model by adding a path (e.g., residual correlation), and model fit was reevaluated. If possible, this process was repeated to find the best model fit.

To investigate construct and criterion-related validity, research questions 2 and 3, multilevel regression analyses were conducted, accounting for clustering of children within both classrooms and daycare centers. For the second research question age and gender were used as predictors while controlling for the number of children during the observation, whether the teacher was part of the activity, and whether it was a teacher-directed activity. Regarding the third research question the inCLASS domain scores were used to predict children's social-emotional skills, the Empathy index and the Self-Regulation & Cooperation index, and their language and pre-math skills. We controlled for children's age and gender, the number of children during the observation and if the teacher was part of the activity, and whether it was a teacher-directed activity. We used standardized regression coefficient as measures of effect size with $\beta < 0.10$ indicating a small effect, β of around 0.30 indicating a medium effect and β of around 0.50 indicating a large effect (Kline, 2005).

4. Results

Descriptive statistics and correlations for the Danish and Dutch sample can be found in Tables 4, 5 and 6 respectively. The results in Table 4 show that children's interactions with teachers are characterized by moderate levels of positive engagement and a general lack of negativity. The quality of communication appears low. Overall, for Danish children the quality of peer interactions is low except for peer sociability. Whereas for Dutch children, the quality of peer interactions appears to be in the midrange. Again, negativity does not occur frequently in interactions with peers either. Children's classroom behavior and task engagement is considered of mid to high quality and the level of self-reliance is solidly in the midrange. The ICC's indicate the level of shared variance at the classroom level. For Danish children, it appears that the domains of teacher interactions and task orientation show moderate levels of shared classroom variance (34 % and 36 %, respectively). For the Dutch children, the domains of Peer Interactions and Behavior Control revealed classroom variance of 19 % and 20 %, respectively.

Tables 5 and 6 present the intercorrelations between the dimensions and domains for both countries separately. The findings illustrate that for Danish children, both higher quality interactions with teachers and peers is associated with higher task orientation. For Dutch children the same pattern is found, but in addition to that the quality of teacher interactions is positively associated with peer interactions.

Table 4
Descriptive statistics of the inCLASS Toddler dimensions on child level.

Dimensions	DK				NL			
	M	SD	Range	ICC	M	SD	Range	ICC
Pos. Eng. w. Teacher	3.92	1.06	2.00–6.70	0.40	3.84	0.89	2.00–6.00	0.09
Teacher Communication	2.94	1.04	1.00–6.70	0.24	3.35	1.06	1.30–6.00	0.08
Teacher Conflict*	6.86	0.32	5.30–7.00	0.03	6.90	0.23	5.70–7.00	0.00
Peer Sociability	2.99	0.87	1.30–5.30	0.22	3.66	0.78	2.00–5.30	0.15
Peer Communication	1.81	0.69	1.00–4.30	0.05	2.59	0.90	1.00–4.70	0.25
Peer Assertiveness	1.86	0.81	1.00–4.30	0.10	2.13	0.87	1.00–5.00	0.11
Peer Conflict*	6.82	0.33	5.00–7.00	0.12	6.79	0.32	5.50–7.00	0.02
Task Engagement	4.93	0.84	1.70–6.70	0.27	4.87	0.77	3.00–6.30	0.03
Self-Reliance	4.10	1.20	1.70–6.70	0.42	3.76	0.95	1.00–5.70	0.02
Behavior Control	6.41	0.61	4.70–7.00	0.08	5.47	0.84	3.00–7.00	0.25
Domains								
Teacher Interactions	3.43	0.99	1.50–6.70	0.34	3.59	0.93	1.70–5.80	0.08
Peer Interactions	2.22	0.73	1.20–4.70	0.15	2.79	0.80	1.30–4.70	0.19
Task Orientation	4.51	0.92	1.70–6.50	0.36	4.32	0.78	2.00–5.80	0.01
Behavior Control*	6.69	0.33	5.10–7.00	0.07	6.39	0.38	4.80–7.00	0.20

Note. Danish sample $n = 97$, Dutch sample $n = 114$. *Includes reversed scales. Aggregated classroom level scores used for the ICCs on classroom level.

Table 5
Bivariate Pearson correlations for the Danish inCLASS Toddler scores.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Dimensions														
1. Positive Eng. w. Teacher	1.00													
2. Teacher Communication	0.78***	1.00												
3. Teacher Conflict ^a	0.01	0.10	1.00											
4. Peer Sociability	-0.08	0.11	0.06	1.00										
5. Peer Communication	-0.23	0.06	0.06	0.74***	1.00									
6. Peer Assertiveness	-0.08	0.13	0.01	0.78***	0.75***	1.00								
7. Peer Conflict ^a	0.08	0.06	0.41**	-0.14	-0.20	-0.20	1.00							
8. Task Engagement	0.23	0.42***	0.17	0.33*	0.30	0.34*	0.05	1.00						
9. Self-Reliance	0.25	0.42***	0.03	0.34*	0.37*	0.48***	-0.05	0.62***	1.00					
10. Behavior Control	-0.06	-0.06	0.45***	-0.19	-0.24	-0.29	0.44***	0.08	-0.30	1.00				
Domains														
11. Teacher Interactions											1.00			
12. Peer Interactions											-0.01	1.00		
13. Task Orientation											0.39***	0.44***	1.00	
14. Behavior Control ^a											0.00	-0.21	-0.07	1.00

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. Bonferoni correction.

^a Includes reversed scales.

Table 6
Bivariate Pearson correlations for the Dutch inCLASS Toddler scores.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Dimensions														
1. Pos. Eng. w. Teacher														
2. Teacher Communication	0.80***	1.00												
3. Teacher Conflict ^a	0.15	0.08	1.00											
4. Peer Sociability	0.56***	0.59***	0.10	1.00										
5. Peer Communication	0.40***	0.51***	0.02	0.82***	1.00									
6. Peer Assertiveness	0.36**	0.55***	-0.09	0.68***	0.81***	1.00								
7. Peer Conflict ^a	-0.03	-0.045	0.39**	-0.09	-0.28	-0.44***	1.00							
8. Task Engagement	0.54***	0.57***	0.18	0.48***	0.40***	0.37***	0.11	1.00						
9. Self-Reliance	0.50***	0.58***	0.08	0.40***	0.40***	0.43***	-0.09	0.66***	1.00					
10. Behavior Control	0.26	0.14	0.47***	0.07	-0.09	-0.31*	0.36**	0.15	0.08	1.00				
Domains														
11. Teacher Interactions											1.00			
12. Peer Interactions											0.57***	1.00		
13. Task Orientation											0.64***	0.49***	1.00	
14. Behavior Control ^a											0.16	-0.18	0.10	1.00

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. Bonferoni correction.

^a Includes reversed scales.

4.1. Factor structure of the inCLASS Toddler

In our first research question we examined the four-domain structure of the inCLASS Toddler with two CFAs separately for the Danish and Dutch data. See Figs. 1 and 2 for the graphical representations of the final factor models. When fitting the factor structure for the Danish data, the model results showed a negative residual for the dimension Teacher Communication, indicating a Heywood case (Geiser, 2013). Since the residual variance was nonsignificant, we constrained it to zero, but model fit values still fell outside the recommended range (RMSEA = 0.125; CFI = 0.895; TLI = 0.843). Based on modification indices, a residual correlation was added between the dimensions Behavior Control and Self-Reliance, but the model fit was still below advised ranges (RMSEA = 0.102; CFI = 0.932; TLI = 0.895). A final modification was made by adding an extra residual correlation between the dimensions Peer Communication and Positive Teacher Engagement, resulting a marginal to adequate model fit (RMSEA = 0.075; CFI = 0.964; TLI = 0.943). Hereafter, the modification indices did not show any more possibilities of adding residual correlations in the model.

Fitting the 4-domain structure of the inCLASS Toddler in the Dutch sample at first resulted in a poor model fit (RMSEA = 0.168; CFI = 0.858; TLI = 0.779). The model fit improved after adding a residual correlation of Behavior Control and Peer Assertiveness (RMSEA = 0.156; CFI = 0.880; TLI = 0.807), but further improved to a marginal model fit when the residual correlations of Peer Assertiveness and Peer Conflict was added (RMSEA = 0.124; CFI = 0.928; TLI = 0.879) and Peer

sociability and Positive Engagement added (RMSEA = 0.106; CFI = 0.949; TLI = 0.911). No further improvements could be made based on the modification indices. Also, a 3-domain structure was fitted, yet this resulted in a poorer fit (RMSEA = 0.118; CFI = 0.934; TLI = 0.891) despite adding five extra residual correlations, amongst it between Peer Conflict and Teacher Conflict and Teacher Conflict and Behavior Control, hinting at a fourth domain.

4.2. Construct validity: age and gender differences

For the second research question, we examined the inCLASS Toddler's sensitivity to gender and age differences with multilevel regression analyses. Within the Danish sample, associations with age were found for the Peer Interactions ($b = 0.070, SE = 0.015, p \leq 0.001; \beta = 0.097$) and Task orientation domain ($b = 0.046, SE = 0.020, p = 0.023; \beta = 0.050$). Thus, older children scored higher on their interactions with peers and showed higher levels of task orientation, showing small effect sizes. For the Dutch sample, age was positively related with the Teacher Interactions ($b = 0.051, SE = 0.016, p = 0.002; \beta = 0.039$) and Behavior Control domain ($b = 0.015, SE = 0.007, p = 0.025; \beta = 0.027$), showing that older children have higher quality interactions with the teacher and show more adaptive classroom behavior and less conflicts (conflict scales are reversed), though the effect sizes were small. No associations with gender were found in either sample.

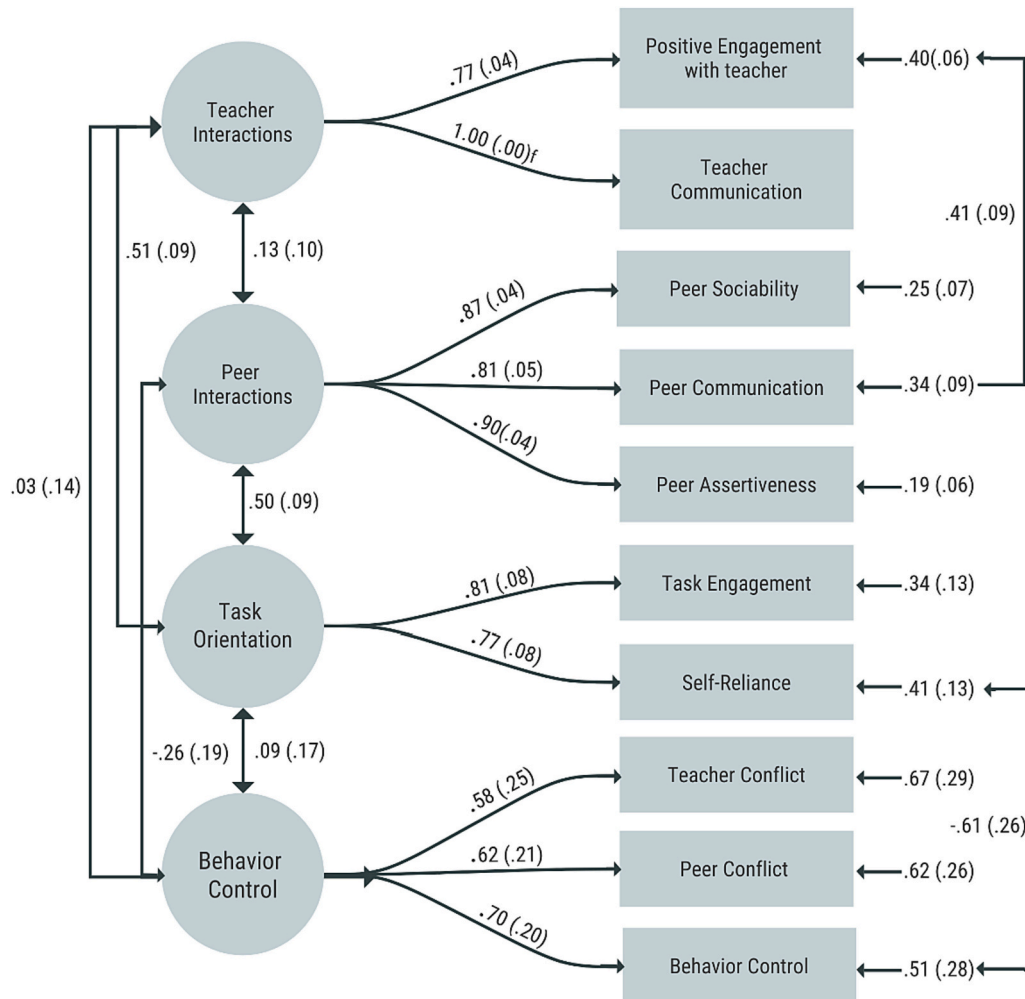


Fig. 1. The final four-domain model for the Danish sample with standardized parameters. The superscript *f* indicates a factor loading fixed to 1 because of a negative residual. Standard errors are given in parentheses. Because of standardization, latent factor variances were 1.00.

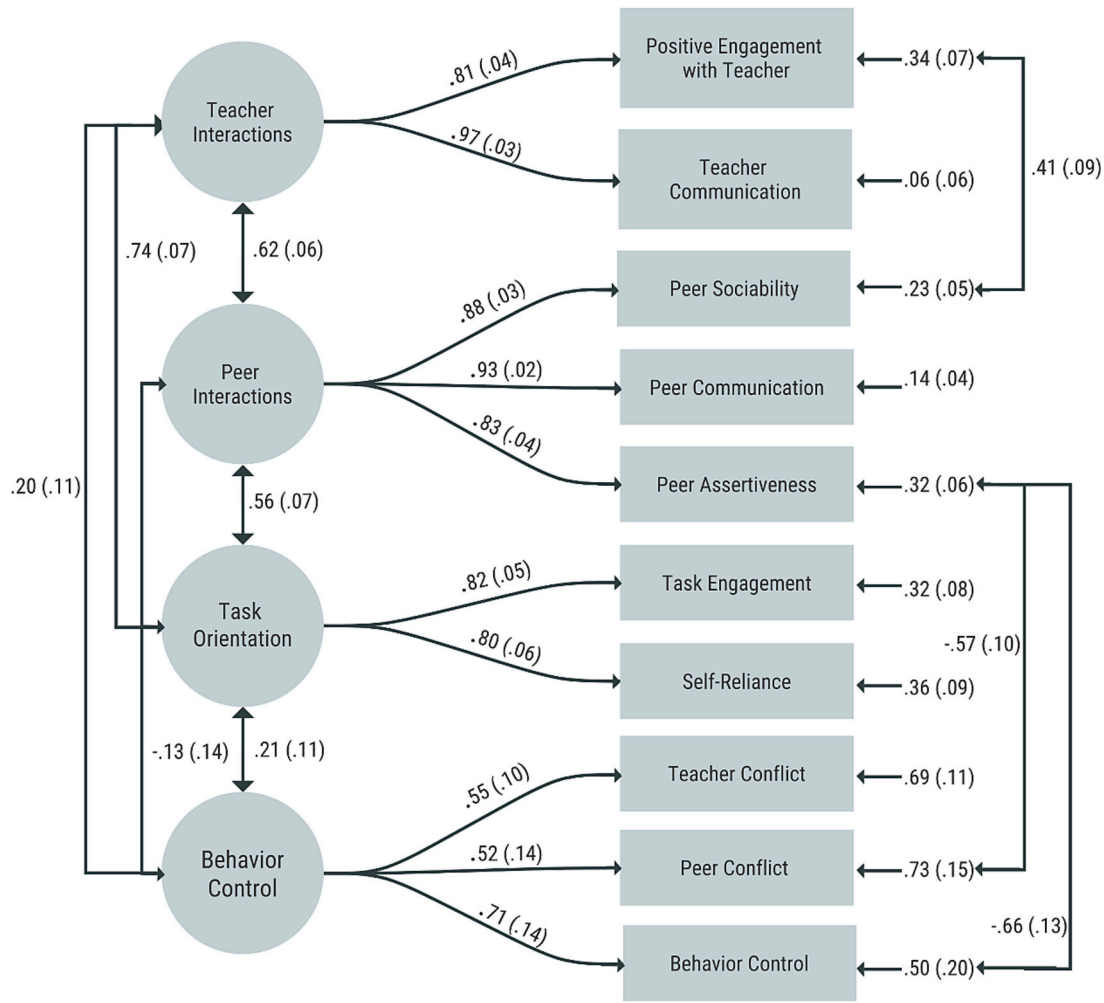


Fig. 2. The final four-domain model for the Dutch sample with standardized parameters. Standard errors are given in parentheses. Because of standardization, latent factor variances were 1.00.

4.3. Criterion validity: relations with social-emotional, language and math development

Multilevel regression analyses were conducted with the Danish inCLASS Toddler factor scores to predict children's language (vocabulary and language use), mathematics (numeracy and math language), and social-emotional skills (empathy and self-regulation and cooperation; see Sjoe et al., 2017), see Tables 7, 8 and 9, revealing some significant associations. Higher levels of positive teacher engagement were related to better vocabulary and math language with small to medium effect sizes. Children's observed behavioral control was positively related to their self-regulation and cooperation skills, revealing a large effect size. A marginally significant effect for the Peer Interaction domain indicated that higher quality peer interactions were associated with better math skills, with a small to medium effect size. Lastly, marginally significant effects with small to medium effect sizes were found for Task Orientation, showing negative associations between toddlers' task related behavior and their empathy and self-regulation and cooperation skills as reported by teachers.

5. Discussion

Through interactions with teachers, peers and the classroom environment, toddlers are provided with varying opportunities to develop in areas such as language, early math, and social-emotional skills. The current study has added to the knowledge base of capturing these crucial

toddler interactions, by providing a first validation of the newly developed inCLASS Toddler across childcare settings in Denmark and the Netherlands. The structural, construct and criterion validity of the inCLASS toddler were investigated.

In the first research question we tested the four-domain factor structure, like the inCLASS Pre-K (Downer et al., 2010), and have found an adequate model fit for the Danish childcare context, whereas for the Dutch context a marginal fit was found. Despite the lack of a good fit, correlations in the Dutch sample hinted towards a 4-domain structure in the data. First, there were positive correlations between Teacher Conflict, Peer Conflict and Behavior Control suggesting a fourth factor. Second, when fitting an alternative 3-domain structure on the Dutch data, modification indices suggested a better model fit specifically when residual correlations between the Peer Conflict and Teacher Conflict dimensions, and Teacher Conflict and Behavior Control were added. Taken together, these results seem to suggest that both in the Danish and Dutch childcare context, the inCLASS Toddler showed potential in capturing individual toddlers' interactions across the earlier four established domains: Teacher interactions, Peer interactions, Task orientation, and Behavior control. This is in line with another Danish study in which the Pre-K version was used (Slot & Bleses, 2018) and other Pre-K studies in Austria, Germany, and the U.S. (Booren et al., 2012; Downer et al., 2018; Schmidt & Embacher, 2021; Vitiello, Moas, et al., 2012; Von Suchodoletz et al., 2015). Although the model fit of the 4-domain-structure was not ideal, it did provide the best fit to the present data. Further research is warranted to corroborate and strengthen

Table 7
Regression coefficients (SE) for criterion validity with Danish toddlers' socio-emotional skills.

	Empathy			
	Teacher interactions	Peer interactions	Task orientation	Behavior control
Gender	2.35 (1.62)	2.29 (1.64)	2.50 (1.59)	2.26 (1.62)
Age	0.13 (0.19)	0.18 (0.21)	0.24 (0.19)	0.18 (0.19)
Teacher-child ratio	-4.43 (3.94)	-4.07 (3.86)	-3.24 (3.78)	-4.05 (3.83)
Teacher part of activity	3.28 (3.02)	3.34 (3.17)	3.06 (2.90)	3.46 (2.92)
Teacher directed activity	0.56 (4.59)	0.81 (4.61)	1.54 (4.52)	1.34 (4.62)
inCLASS Toddler domain	0.52 (0.96)	-0.40 (1.39)	-1.81+ (1.01)	2.76 (2.91)
Effect size (β)	0.07	-0.05	-0.25	0.36

	Self-regulation & cooperation			
	Teacher interactions	Peer interactions	Task orientation	Behavior control
Gender	-0.24 (1.03)	-0.33 (1.04)	-0.14 (1.02)	-0.35 (0.99)
Age	-0.13 (0.12)	-0.09 (0.13)	-0.09 (0.12)	-0.11 (0.12)
Teacher-child ratio	-0.58 (2.51)	-1.03 (2.47)	-0.38 (2.44)	-1.12 (2.37)
Teacher part of activity	0.79 (1.97)	-0.29 (2.07)	0.04 (1.90)	0.17 (1.87)
Teacher directed activity	-6.22* (2.96)	-6.03* (2.96)	-5.82 (2.93)*	-5.60+ (2.89)
inCLASS Toddler domain	-0.37 (0.63)	-0.86 (0.90)	-1.09+ (0.65)	3.99* (1.81)
Effect size (β)	-0.07	-0.17	-0.22	0.79

Note. ***p < 0.001, **p < 0.01, *p < 0.05, +p < 0.10. Observations for each model: n = 75, number of groups: n = 17, number of childcares n = 23.

the current findings, though.

Interestingly, the inter-correlations showed some differential relations in Denmark and the Netherlands, which may point to cultural differences in pedagogical practices despite the same 4-factor structure in both childcare contexts. In the Dutch context the quality of children's interactions with the teacher was positively associated with their quality of peer interactions, such that children engaging in more positive interactions with the teacher(s) also showed more positive interactions with peers, whereas in the Danish context these were unrelated. Prior work in Danish preschools with the inCLASS Pre-K even revealed negative relations between children's interactions with teachers and their interactions with peers, suggesting that children are either engaged in teacher interactions or in peer interactions but not simultaneously (Slot & Bleses, 2018). The Dutch findings showed that higher quality of teacher interactions coincides with higher peer interactions, suggesting a different pattern of interactions in childcare. Indeed, the Dutch observations showed overall higher presence and engagement of teachers in children's activities and play (95 %) as compared to the Danish findings where children were also observed to be without immediate presence of the teacher in a third of the observations (33 %). On a related note, the Dutch children appeared to be more involved in organized educational and creative activities compared to the Danish children who spent most of their time in free play inside or outdoors. Although the difference in coding procedures (live coding and video coding) plays a role, it also partly reflects differences in provided activities. For instance, another Dutch study using a time-sampling method showed that children on average spent 15 % of their time in free play, almost 12 % in

Table 8
Regression coefficients (SE) for criterion validity with Danish toddlers' pre-language skills.

	Vocabulary			
	Teacher interactions	Peer interactions	Task orientation	Behavior control
Gender	-0.57 (2.74)	-0.01 (2.87)	-0.57 (2.81)	-0.57 (2.84)
Age	3.38*** (0.35)	3.32*** (0.40)	3.37*** (0.37)	3.52*** (0.36)
Teacher-child ratio	-5.31 (6.29)	-1.70 (6.29)	-3.29 (6.23)	-2.48 (6.29)
Teacher part of activity	-5.42 (4.75)	-0.79 (4.94)	-2.45 (4.57)	-2.87 (4.66)
Teacher directed activity	1.58 (7.38)	2.42 (7.52)	1.20 (7.60)	3.47 (7.62)
inCLASS Toddler domain	3.21* (1.60)	2.53 (2.35)	2.50 (1.84)	0.93 (4.89)
Effect size (β)	0.16	0.13	0.12	0.05

	Language use			
	Teacher interactions	Peer interactions	Task orientation	Behavior control
Gender	0.30 (0.51)	0.32 (0.53)	0.32 (0.51)	0.32 (0.52)
Age	0.40*** (0.07)	0.40*** (0.08)	0.43*** (0.07)	0.40*** (0.07)
Teacher-child ratio	-0.60 (1.21)	-0.47 (1.24)	-0.42 (1.21)	-0.49 (1.23)
Teacher part of activity	0.09 (1.04)	0.58 (1.09)	0.52 (0.99)	0.60 (1.01)
Teacher directed activity	-0.09 (1.46)	0.06 (1.48)	0.45 (1.49)	0.02 (1.49)
inCLASS Toddler domain	0.43 (0.34)	0.03 (0.47)	-0.53 (0.37)	-0.26 (0.93)
Effect size (β)	0.14	0.01	-0.17	-0.08

Note. ***p < 0.001, **p < 0.01, *p < 0.05, +p < 0.10. Observations for each model: n = 69, number of groups: n = 17, number of childcares n = 23.

educational activities and almost 9 % in creative activities (De Haan et al., 2011). Thus, this reveals a more balanced program of play and activities throughout a morning as compared to the Danish context. Thus, the inCLASS illustrates the differences in individual children's experiences in childcare that an overall classroom quality measure may not capture. In fact, overall classroom quality in Danish childcare, as assessed with the CLASS Toddler, showed higher levels of Emotional Support (M = 5.72, SD = 0.54; Bleses et al., 2020) compared to the Dutch childcare centers (M = 5.33, SD = 0.45; Wyslowska & Slot, 2020) and the opposite pattern for Engaged Support for Learning (M = 2.98, SD = 0.59 in Danish centers and M = 3.57, SD = 0.61 in Dutch centers respectively). Taken together, it is promising that the inCLASS evaluates children's experiences in a more detailed way, capturing relevant cultural and contextual variation in classrooms, which adds to the body of knowledge on classroom quality from the perspective of children. It would be interesting for future work to investigate differences in activity settings.

Observations from the perspective of individual children can have important implications for practice, as it may help teachers to obtain a better understanding of how experiences of individual children may vary in the classroom. Through the interaction with their environment children learn to express and enact their competences, reflecting the transactional nature. Some children may seek out more active opportunities for exploration and interaction with their environment and, as

Table 9
Regression coefficients (SE) for criterion validity with Danish toddlers' pre-math skills.

	Numeracy			
	Teacher interactions	Peer interactions	Task orientation	Behavior control
Gender	0.91 (0.97)	1.15 (0.96)	0.91 (0.97)	0.88 (0.98)
Age	1.05***(0.13)	0.94***(0.14)	1.06***(0.13)	1.06***(0.12)
Teacher-child ratio	-0.51 (2.28)	0.03 (2.20)	-0.51 (2.23)	-0.53 (2.23)
Teacher part of activity	-1.02 (1.75)	0.18 (1.80)	-1.08 (1.69)	-1.12 (1.69)
Teacher directed activity	-1.26 (2.67)	-1.64 (2.61)	-1.19 (2.70)	-1.10 (2.68)
inCLASS Toddler domain	-0.03 (0.58)	1.48 ⁺ (0.83)	-0.12 (0.65)	0.79 (1.71)
Effect size (β)	-0.01	0.23	-0.02	0.12

	Math language			
	Teacher interactions	Peer interactions	Task orientation	Behavior control
Gender	-0.71 (1.47)	-0.43 (1.63)	-0.43 (1.62)	-0.77 (1.58)
Age	2.07***(0.20)	2.06***(0.24)	2.17***(0.22)	2.14***(0.21)
Teacher-child ratio	0.64 (3.59)	0.98 (3.89)	0.86 (3.86)	1.11 (3.78)
Teacher part of activity	-0.94 (3.20)	2.61 (3.41)	1.74 (3.09)	1.01 (3.14)
Teacher directed activity	-4.43 (4.32)	-3.33 (4.62)	-2.53 (4.67)	-2.73 (4.55)
inCLASS Toddler domain	2.44* (1.03)	0.69 (1.48)	-1.12 (1.13)	4.03 (2.85)
Effect size (β)	0.20	0.06	-0.09	0.33
Observations	75	75	75	75
Number of groups	17	17	17	17
Number of childcares	23	23	23	23

Note. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$. Observations for each model: $n = 69$, number of groups: $n = 17$, number of childcares $n = 23$.

such, elicit more attention and interaction from teachers and peers in that environment. The current findings indicated substantial variance of children's quality of peer interactions, task orientation and behavioral control at the classroom level, which seems to point to the fact that some classrooms, and thus their teachers, provide children with more opportunities to develop peer interaction skills, adaptive classroom behavior and engagement in play and activities. As such, focusing on individual children may support teachers in having a better understanding of how their children interact in the classroom, which may help them to better plan and tailor their daily activities to meet diverse needs of children.

For the second research question we examined the sensitivity of the inCLASS Toddler in distinguishing individual differences in terms of children's gender and age. Contrary to our expectations, there were no differences between boys and girls both in the combined and in the separate Danish and Dutch samples. In a Danish study with preschoolers, small differences were found, such that boys showed higher levels of conflicts than girls, as was the case in studies of preschoolers (Downer et al., 2018; Slot & Bleses, 2018; Von Suchodoletz et al., 2015). This may suggest that differential patterns in children's interactions develop between toddlerhood and preschool age. A gender socialization perspective is often used to explain why boys tend to exhibit higher levels of conflict than girls (e.g., Ewing & Taylor, 2009; Hamre et al., 2007). Consequently, this developmental period provides an interesting window of opportunity for supporting the development of healthy and positive interactions, especially for boys. Regarding age, as anticipated the quality of Danish toddlers' interaction with peers increased with age, which was not the case for the Dutch toddlers. The Dutch toddlers were on average older in age and showed less variation, which may explain the null associations regarding peer interactions. The results for the Danish toddlers could illustrate the increase in peer orientation and peer interactions with age, following the natural course of children's development. Another finding relates to the quality of toddlers' interactions with the teacher, which increased with age in Dutch toddlers. This may reflect the increased communicative skills that enabled toddlers to engage in more qualitative exchanges with the teacher. Note that this domain includes teacher communication, which deliberately aimed to capture both non-verbal communicative attempts and verbal ones. However, the scale is focused on the child's initiative and perseverance in (non)verbal interactions. Lower scores may reflect fewer child

initiatives, but perhaps also fewer initiatives from the part of the teacher. Perhaps teachers have different interaction styles when interacting with children who are less proficient in language. In future studies, it would be interesting to address this reciprocity in more detail to further entangle the origin of this finding. Another explanation could be related to the heterogeneity of the Dutch sample. The sample was more diverse in comparison to the Danish sample and consisted of children from low SES backgrounds and/or speaking another home language. Consequently, some of these children are known to have a language delay, which may mean they feel less comfortable in communicating (non)verbally with teachers or teachers may interact differently with these children. More work in diverse samples is needed to explore explanations for differential relations. For Dutch toddlers, behavior control was positively associated with age, such that older children showed less conflictual interactions and more appropriate classroom behavior. For Danish toddler, task orientation increased with age, which is in line with previous work in preschoolers in the U.S. (Downer et al., 2018; Vitiello, Moas, et al., 2012; Williford et al., 2013), though prior work in Denmark showed the opposite (Slot & Bleses, 2018). More work is needed to better understand the developmental patterns of young children's task orientation and behavioral control in the classroom setting.

Lastly, in addressing the third research question, we found small to large effect sizes for the associations between Danish toddlers' quality of interactions and their social-emotional, language and math skills. Higher quality interactions with the teacher were associated with better language and math skills, which is in line with previous work (Thompson, 2006) also with the inCLASS Pre-K (Sabol et al., 2018). Also, higher quality of peer interactions showed marginal positive relations with children's math skills. It suggests that both the teacher and peers can play a vital role in the development of toddlers. Yet children's task orientation showed marginally significant negative associations with their social-emotional skills based on teacher ratings, with small to medium effect sizes, which is not in line with evidence from the inCLASS Pre-K (Kim et al., 2019; Williford et al., 2013). This may point to developments that are unique to toddlerhood. Self-regulation skills develop rapidly during toddlerhood but at the same time there is a strong increase in toddlers' motivation and will to demonstrate independence and self-direction (Bronson, 2000). Perhaps a higher score on task orientation reflects a stronger tendency to express self-reliance

which may be slightly at the expense of being able to take other children's perspectives or desires into account and act accordingly. Also, toddlers who may be strong-willed and independent, also referred to as the 'terrible twos', may be viewed as less compliant by teachers. Toddlers are still very much reliant on adults to guide and help them to balance their need for autonomy and independence whilst maintaining satisfactory and positive relations with others (Bronson, 2000). It would be interesting for future research to further explore this. Finally, a large effect pointed towards the finding that children who displayed more classroom appropriate behavior and less conflicts with the teacher and peers, were reported to have better self-regulation and cooperation skills, in line with inCLASS Pre-K findings (Kim et al., 2019). This Behavior Control domain (in other studies often reversely coded and labeled as Conflictual Interactions or Negative Engagement) has been criticized in some studies due to a low variability (e.g., Schmidt & Embacher, 2021) and even excluded from analyses altogether (e.g., Guedes et al., 2020). However, the current study provides some support that even with limited variance, this variance may capture important aspects of children's competences, in line with prior work with the inCLASS Pre-K (Slot & Bleses, 2018).

Despite its strengths, this exploratory, cross-country validation study is subject to several limitations. First, although the sample size was substantive for an observational study, the relatively small sample size for a confirmatory factor analysis may still have interfered with establishing a good fit in the models. Future research could look into the possibility of investigating the extent to which the factor structure is comparable across contexts (i.e., establish measurement invariance) in order to compare inCLASS scores directly across countries. Second, the age range for the Dutch toddlers was different from the age range of the Danish toddlers and showed a more skewed distribution, which may explain the poorer model fit for the Dutch toddlers. However, we explored some different options for the Dutch sample only, such as excluding the 20 % oldest children and separating the half-day from the full-day programs, but both did not improve the model fit. It could be that the heterogeneity of the sample regarding the age and socioeconomic and linguistic background of the children, in combination with the small sample size resulted in a poorer model fit compared to the Danish sample. Relatedly, the inter-observer reliability was poorer for some dimensions in the Dutch sample, such as Peer assertiveness and Behavior control. It seems that the difficulty is especially related to the child's positive and adaptive classroom behavior whether or not directed to peers. All coders passed a reliability test prior to coding (with an average score of 85 %). In another study (Blinded for review, 2017, 2018) inCLASS coding was also done based on videos and this study revealed an ICC of 0.82 (based on the same inCLASS training). One possible explanation for the lower reliability in the current study could perhaps be the relatively high socioeconomic and linguistic diversity in the sample, also compared to the Danish sample. Thus, further research with the inCLASS Toddler in larger and more diverse samples is warranted, which has as only recently been explored with the inCLASS Pre-K (Bohmann et al., 2019). Also, diversity in the classroom can be addressed more strongly in training and preparing for coding. Another limitation concerns the different approaches to the inCLASS coding in Denmark and the Netherlands. In Denmark a live coding procedure was used, whereas in the Netherlands coding was done from video. In this case, both methods worked well. An important goal when collecting data was to observe the natural course of the day while minimizing the burden for teachers. For the Danish settings the use of live coding was deemed the most appropriate in this case. In the Netherlands, because of a more structured program of play and organized activities during the day, the use of videos was suitable and allowed for the use of multiple observation measures, such as the CLASS toddler (not the focus of the current paper, but addressed elsewhere; Wysłowska & Slot, 2020). Lastly, it is a limitation that the criterion validity was only examined for

the Danish inCLASS Toddler scores. Also, the child outcomes are teacher-reported, and we cannot rule out a potential bias in teachers' evaluation of children's self-regulation and cooperation, empathy and (math) language skills for children they are more closely connected to and share more interactions with. However, it is not uncommon to rely on teacher reports. For instance, another study that studied relations between the inCLASS Pre-K and children's social and behavioral skills also relied on teacher reports (Kim et al., 2019). In all, this validation study should be regarded as exploratory, until future research can obtain larger samples and apply sampling methods in which the strata are more refined in terms of age and other background factors and other child outcome data are used.

To conclude, this exploratory validity study adds to the early childhood education and care practice and research field in several ways. First, we delivered a new observation instrument specifically aimed at individual toddlers and their in-depth, situated interactions with not only the teachers, but also with peers and within tasks and activities. Such a direct observational measure provides additional information on how children experience quality and complements classroom-level measures to assess process quality in large-scale cohort and randomized control trial studies. Changing the lens through which you evaluate classroom practices can be a promising way to improve practice. Similar to how the classroom-level process quality measure CLASS has shown to be useful in enhancing teachers' professional development (Downer et al., 2009) and improving the quality of interactions (Egert et al., 2020), the inCLASS has also shown to be useful (Downer et al., 2018). Taking advantage of a systematic observation method could stimulate teachers to self-reflect and be sensitive to a broad spectrum of toddlers' skills and interactions, and ways to develop these. Social interactions within toddlerhood are often overlooked in early childhood education and care research and interventions on social-emotional competences (O'Flaherty et al., 2019). Second, the current results support the applicability of the inCLASS Toddler in different contexts, which is encouraging and should be validated further by larger, more parallel samples. Future research can address larger and more diverse samples to increase the evidence on the psychometric properties of the inCLASS Toddler, especially regarding the predictive validity. Like the inCLASS Pre-K, which has been used in widespread research, this tool for toddlers is useful given the importance of development in toddlerhood for school readiness and future success. Taken together, this paper highlights the inCLASS Toddler as a promising observation instrument for measuring individual toddlers' interactions with teachers and peers across and within tasks and activities.

CRediT authorship contribution statement

Pauline Slot: Conceptualization, writing – original draft preparation; Marinka Willemsen: Conceptualization, writing – original draft preparation; Dorthe Bleses: Conceptualization, methodology, writing – project administration, funding acquisition; Jason Downer: Conceptualization, writing – reviewing and editing.

Declaration of competing interest

The authors declare there is no competing interest at stake.

Acknowledgements

The Danish research was supported from a grant by the Ministry of Children and Education and TrygFonden (project number 21865). The Dutch work was part of the research project QualityMatters (2016–2019) supported by the FCT — Fundação para a Ciência e a Tecnologia under Grant [PTDC/MHC-CED/5913/2014]. We wish to thank all participating municipalities, teachers, and children.

Appendix A. Overview of adaptations from the Pre-K version to the Toddler version

Dimensions	Changes indicators	Changes behavioral markers
Positive Engagement with Teacher		<i>Attunement</i> Cooperation deleted <i>Proximity Seeking</i> Conversation → Communication Emotional “refueling” added
Teacher Communication	Sustains Conversation → Initiates Communication	<i>Initiates Communication</i> Leads → Seek contact <i>Communicative Interactions</i> Contingency and Topic prolonged deleted
Teacher Conflict	Aggression → Negativity	<i>Negative Affect:</i> Whining and Resist contact added <i>Attention-seeking</i> Complaining → Dissatisfaction. <i>Noncompliance</i> Argumentative → Uncooperative <i>Social Relationships</i> Conversation → Communication. <i>Peer Acceptance</i> Friendship deleted
Peer Sociability	Cooperation → Social Relationships Popularity → Peer Acceptance	<i>Social Relationships</i> Conversation → Communication. <i>Peer Acceptance</i> Friendship deleted
Peer Communication	Communicative Interactions → Communicative interactions	<i>Initiates Communication</i> Leads → Seek contact Imitate added <i>Communicative Interactions</i> Contingency and Topic prolonged deleted
Peer Assertiveness		<i>Initiations</i> Joining groups → Joining peers Conversations → Communication <i>Leadership</i> Organizes play → Initiates play Teaches peers deleted
Peer Conflict	Aggression and Confrontation merged → Negativity	<i>Negativity</i> Relational → Uncooperative <i>Negative Affect</i> Whining and Resist contact added <i>Attention-seeking</i> Complaining → Dissatisfaction <i>Personal initiative</i> Inquisitive and Linking concepts deleted Exploration has been added. Walks has been deleted
Self-Reliance		
Behavior Control	Compliance added	

Note. No adaptations were made for the dimension Task Engagement.

References

- Bleses, D., Jensen, P., Højen, A., & Dale, P. S. (2018). An educator-administered measure of language development in young children. *Infant Behavior and Development, 52*, 104–113. <https://doi.org/10.1016/j.infbeh.2018.06.002>
- Bleses, D., Jensen, P., Slot, P., & Justice, L. (2020). Low-cost teacher-implemented intervention improves toddlers’ language and math skills. *Early Childhood Research Quarterly, 53*, 64–76. <https://doi.org/10.1016/j.ecresq.2020.03.001>
- Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T., & Basbøll, H. (2008). The Danish Communicative Development Inventories: Validity and main developmental trends. *Journal of Child Language, 35*, 651–669. <https://doi.org/10.1017/S0305000907008574>
- Bohlmann, N. L., & Downer, J. T. (2016). Self-regulation and task engagement as predictors of emergent language and literacy skills. *Early Education and Development, 27*(1), 18–37. <https://doi.org/10.1080/10409289.2015.1046784>
- Bohlmann, N. L., Downer, J. T., Williford, A. P., Maier, M. F., Booren, L. M., & Howes, C. (2019). Observing children’s engagement: Examining factorial validity of the inCLASS across demographic groups. *Journal of Applied Developmental Psychology, 60*, 166–176 (<https://doi.org/20.1016/j.appdev.2018.08.007>).
- Booren, L. M., Downer, J. T., & Vitiello, V. E. (2012). Observations of children’s interactions with teachers, peers, and tasks across preschool classroom activity settings. *Early Education & Development, 23*, 517–538. <https://doi.org/10.1080/10409289.2010.548767>
- Broekhuizen, M. L., Slot, P. L., Dubas, J. S., & van Aken, M. A. G. (2016). Teacher’s emotional and behavioral support and preschoolers’ self-regulation: Relations with social and emotional skills during play. *Early Education and Development, 28*, 135–153. <https://doi.org/10.1080/10409289.2016.1206458>
- Broekhuizen, M. L., van Aken, M. A. G., Dubas, J. S., Mulder, H., & Leseman, P. P. M. (2015). Individual differences in effects of child care quality: The role of child affective self-regulation and gender. *Infant Behavior and Development, 40*, 216–230. <https://doi.org/10.1016/j.infbeh.2015.06.009>
- Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes. In W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology* (pp. 993–1028). New York, NY: John Wiley.
- Bronson, M. B. (2000). *Self-regulation in early childhood. Nature and nurture* (pp. 86–112, 167–198). New York, NY: The Guilford Press.
- Burchinal, M. R. (2018). Measuring early care and education quality. *Child Development Perspectives, 12*(1), 3–9.
- Cadima, J., Aguiar, C., Guedes, C., Wysłowska, O., Salminen, J., Slot, P., ... Lerkkanen, M. K. (2023). Process quality in toddler classrooms in four European countries. *Early Education and Development, 1–25*. <https://doi.org/10.1080/10409289.2022.2139548>
- Calkins, S. (2007). The emergence of self-regulation: Biological and behavioral control mechanisms supporting toddler competencies. In C. A. Brownell, & C. B. Kopp (Eds.), *Socioemotional development in the toddler years: Transitions and transformations*. New York, NY: The Guilford Press. <https://doi.org/10.5860/choice.45-6470>
- Chen, J., Justice, L. M., Tambyraja, S. R., & Sawyer, B. (2020). Exploring the mechanism through which peer effects operate in preschool classrooms to influence language growth. *Early Childhood Research Quarterly, 53*, 1–10. <https://doi.org/10.1016/j.ecresq.2020.02.002>
- Connor, C. M., Adams, A., Zargar, E., Wood, T. S., Hernandez, B. E., & Vandell, D. L. (2020). Observing individual children in early childhood classrooms using Optimizing Learning Opportunities for Students (OLOS): A feasibility study. *Early*

- Childhood Research Quarterly*, 52, 74–89. <https://doi.org/10.1016/j.ecresq.2019.10.001>
- Côté, S. M., Japel, C., Séguin, J. R., Mongeau, C., Xu, Q., & Tremblay, R. E. (2013). Child care quality and cognitive development: Trajectories leading to better preacademic skills. *Child Development*, 84, 752–766. <https://doi.org/10.1111/cdev.12007>
- De Haan, A., Leseman, P., & Elbers, E. (2011). Pilot gemengde groepen 2007-2010. In *Onderzoeksrapportage 2011*. Utrecht: Universiteit Utrecht.
- Downer, J., LoCasale-Crouch, J., Hamre, B., & Pianta, R. (2009). Teacher characteristics associated with responsiveness and exposure to consultation and online professional development resources. *Early Education & Development*, 20, 431–455. <https://doi.org/10.1080/10409280802688626>
- Downer, J., Williford, A., Bultoksy-Shearer, R. J., Vitiello, V. E., Bouza, J., Reilly, S., & Lhospatial, A. (2018). Using data-driven, video-based early childhood consultation with teachers to reduce children's challenging behaviors and improve engagement in preschool classrooms. *School Mental Health*, 10, 226–242. <https://doi.org/10.1007/s12310-017-9237-0>
- Downer, J. T., Booren, L. M., Lima, O. K., Luckner, A. E., & Pianta, R. C. (2010). The individualized classroom assessment scoring system (inCLASS): Preliminary reliability and validity of a system for observing preschoolers' competence in classroom interactions. *Early Childhood Research Quarterly*, 25, 1–16. <https://doi.org/10.1016/j.ecresq.2009.08.004>
- Egert, F., Dederer, V., & Fukkink, R. G. (2020). The impact of in-service professional development on the quality of teacher-child interactions in early education and care: A meta-analysis. *Educational Research Review*, 29, Article 100309. <https://doi.org/10.1016/j.edurev.2019.100309>
- Ewing, A. R., & Taylor, A. R. (2009). The role of child gender and ethnicity in teacher-child relationship quality and children's behavioral adjustment in preschool. *Early Childhood Research Quarterly*, 24(1), 92–105. <https://doi.org/10.1016/j.ecresq.2008.09.002>
- Fabes, R. A., Hanish, L. D., & Martin, C. L. (2003). Young children at play: The role of peers in understanding the effects of child care. *Child Development*, 74, 1039–1043. <https://doi.org/10.1111/1467-8624.00586>
- Fantuzzo, J. W., Bulotsky-Shearer, R., Fusco, R. A., & McWayne, C. (2005). An investigation of preschool classroom behavioral adjustment problems and social-emotional school readiness competencies. *Early Childhood Research Quarterly*, 20(3), 259–275. <https://doi.org/10.1016/j.ecresq.2005.07.001>
- Fenson, L., Marchman, V., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates communicative development inventories. Users guide and technical manual* (2nd ed.). Baltimore: Paul H. Brookes Publishing Co. <https://doi.org/10.1037/111538-000>
- Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, 134, 31–60. <https://doi.org/10.1037/0033-2909.134.1.31>
- Geiser, C. (2013). *Data analysis with Mplus* (English ed.). New York: The Guilford press.
- Gifford-Smith, M. E., & Brownell, C. A. (2003). Childhood peer relationships: Social acceptance, friendships, and peer networks. *Journal of school psychology*, 41(4), 235–284. [https://doi.org/10.1016/S0022-4405\(03\)00048-7](https://doi.org/10.1016/S0022-4405(03)00048-7)
- Guedes, C., Cadima, J., Aguiar, T., Aguiar, C., & Barata, C. (2020). Activity settings in toddler classrooms and quality of group and individual interactions. *Journal of Applied Developmental Psychology*, 67, Article 101100. <https://doi.org/10.1016/j.appdev.2019.101100>
- Hamre, B., Pianta, R., Downer, J., DeCoster, J., Mashburn, A. J., Jones, S. M., ... Hamagami, A. (2007). Teaching through interactions: Testing a developmental framework of teacher effectiveness in over 4,000 classrooms. *Elementary School Journal*, 113, 461–487. <https://doi.org/10.1086/669616>
- Harms, T., Clifford, R. M., & Cryer, D. (2015). *Early Childhood Environment Rating Scale* (Third Edition). New York, NY: Teachers College Press.
- Hoff, E. (2013). Interpreting the early language trajectories of children from low SES and language minority homes: Implications for closing the achievement gap. *Developmental Psychology*, 49, 4–14. <https://doi.org/10.1037/q0027238>
- Kim, H., Cameron, C. E., Kelly, C. A., West, H., Mashburn, A. J., & Grissmer, D. W. (2019). Using an individualized observational measure to understand children's interactions in underserved kindergarten classrooms. *Journal of Psychoeducational Assessment*, 37(8), 935–956. <https://doi.org/10.1177/0734282918819579>
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15, 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Kopp, C. B. (1982). Antecedents of self-regulation: A developmental perspective. *Developmental Psychology*, 18, 199. <https://doi.org/10.1037/0012-1649.18.2.199>
- La Paro, K. M., Hamre, B. K., & Pianta, R. (2012). *Classroom assessment scoring system - Toddler*. Baltimore, Maryland: Paul H. Brookes Publishing Co.
- McClelland, M. M., Morrison, F. J., & Holmes, D. L. (2000). Children at risk for early academic problems: The role of learning-related social skills. *Early Childhood Research Quarterly*, 15, 307–329.
- Melhuish, E., Ereky-Stevens, K., Petrogiannis, K., Ariescu, A., Penderi, E., Rentzou, K., ... Leseman, P. (2015). A review of research on the effects of early childhood education and care (ECEC) upon child development. In *EU FP7 CARE: Curriculum & quality analysis and impact review of European early childhood education and care*. Utrecht University.
- National Scientific Council on the Developing Child. (2009). Young children develop in an environment of relationships. In *Working paper no. 1*. Retrieved from <http://www.developingchild.net>.
- NICHD Early Child Care Research Network. (1998). Early child care and self-control, compliance, and problem behavior at twenty-four and thirty-six months. *Child Development*, 69(4), 1145–1170. <https://doi.org/10.1111/j.1467-8624.1998.tb06165.x>
- NICHD Early Child Care Research Network. (2000). The relation of child care to cognitive and language development. *Child Development*, 71(4), 960–980. <https://doi.org/10.1111/1467-8624.00202>
- NICHD Early Child Care Research Network. (2006). Child-care effect sizes for the NICHD study of early child care and youth development. *American Psychologist*, 61(2), 99–116. <https://doi.org/10.1037/0003-066X.61.2.99>
- OECD. (2023). OECD family data base. Retrieved from <https://www.oecd.org/social/database.htm>.
- O'Flaherty, C., Barton, E. E., Winchester, C., & Domingo, M. (2019). Coaching teachers to promote social interactions with Tddlers. *Journal of Positive Behavior Interventions*, 21, 199–212. <https://doi.org/10.1177/1098300719851794>
- Pianta, R. C., & Stuhlman, M. W. (2004). Teacher-child relationships and children's success in the first years of school. *School Psychology Review*, 33(3), 444–458. <https://doi.org/10.1080/02796015.2004.12086261>
- Rimm-Kaufman, S. E., Curby, T. W., Grimm, K. J., Nathanson, L., & Brock, L. L. (2009). The contribution of children's self-regulation and classroom quality to children's adaptive behaviors in the kindergarten classrooms. *Developmental Psychology*, 45, 958–972. <https://doi.org/10.1037/a0015861>
- Rojas, N. M., & Abenavoli, R. M. (2021). Preschool teacher-child relationships and children's expressive vocabulary skills: The potential mediating role of profiles of children's engagement in the classroom. *Early Childhood Research Quarterly*, 56, 225–235. <https://doi.org/10.1016/j.ecresq.2021.04.005>
- Ruzek, E., Burchinal, M., Farkas, G., & Duncan, G. J. (2014). The quality of toddler child care and cognitive skills at 24 months: Propensity score analysis from the ECLS-B. *Early Childhood Research Quarterly*, 29, 12–21. <https://doi.org/10.1016/j.ecresq.2013.09.002>
- Sabol, T. J., Bohlman, N. L., & Downer, J. T. (2018). Low-income ethnically diverse children's engagement as a predictor of school readiness above preschool classroom quality. *Child Development*, 89, 556–576. <https://doi.org/10.1111/cdev.12832>
- Schmidt, W., & Embacher, E. M. (2021). Examining the factorial validity of the Individualized Classroom Assessment Scoring System in preschools in Austria. *International Journal of Early Years Education*, 1-13. <https://doi.org/10.1080/09669760.2021.1893158>
- Shonkoff, J. P. (2011). Protecting brains, not simply stimulating minds. *Science*, 333, 982–983. <https://doi.org/10.1126/science.1206014>
- Sjoe, N. M., Bleses, D., Jensen, P., Dybdal, L., Kreiner, S., Sehested, K., & Kirkeby, H. (2017). Measurement properties of the Social Emotional Assessment Measure (SEAM) questionnaire using Rasch analysis on data from a representative sample of 0–6-year-olds. *Journal of Psychoeducational Assessment*. <https://doi.org/10.1177/0734282917746250>
- Slot, P. L. (2018). Early childhood education and care in the Netherlands: A shift towards an integrated system aimed at enhancing children's development and learning. In S. Garvis, S. Philipson, & H. Harju-Luukkainen (Eds.), *Early childhood education in the 21st century: International teaching, family and policy perspectives, volume 1 early childhood education in the 21st century: An international perspective*. London/New York: Routledge. <https://doi.org/10.4324/9780203730553>
- Slot, P. L., & Bleses, D. (2018). Individual children's interactions with teachers, peers, and tasks: The applicability of the inclass pre-k in Danish preschools. *Learning and Individual Differences*, 61, 68–76. <https://doi.org/10.1016/j.lindif.2017.11.003>
- Slot, P. L., Bleses, D., & Downer, J. (2016). *The Individualized Classroom Assessment Scoring System (inCLASS) Toddler* [unpublished pilot version for toddlers].
- Slot, P. L., Leseman, P. P. M., Verhagen, J., & Mulder, H. (2015). Associations between structural quality aspects and process quality in Dutch early childhood education and care settings. *Early Childhood Research Quarterly*, 33, 64–76.
- Squires, J. K., Bricker, D., Waddell, M. L., Funk, K., Clifford, J. R., & Hoeselton, R. M. (2014). *Social-Emotional Assessment/Evaluation Measure (SEAMTM)* (Research ed.). Baltimore, MD: Paul H. Brookes. Retrieved from <https://products.brookespublishing.com/Social-Emotional-Assessment/Evaluation-Measure-SEAM-Research-Edition-P717.aspx>.
- Sroufe, L. A. (1996). *Emotional development: The organization of emotional life in the early years*. New York, NY: Cambridge University Press. <https://doi.org/10.1017/cbo9780511527661>
- Thompson, R. A. (1999). Early attachment and later development. In J. Cassidy, & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (pp. 265–286). New York, NY: Guilford Press. <https://doi.org/10.1017/s0021963001226754>
- Thompson, R. A. (2006). Nurturing developing brains, minds, and hearts. In R. Lally, & P. Mangione (Eds.), *Concepts of care: 20 essays on infant/toddler development and learning* (pp. 47–52). Sausalito, CA: WestEd. Retrieved from <https://sedl.faculty.ucdavis.edu/publications/>
- Vach, W., Bleses, D., & Jørgensen, R. N. (2010). Construction of a Danish CDI short form for language screening at the age of 36 months: Methodological considerations and results. *Clinical Linguistics & Phonetics*, 24, 602–621. <https://doi.org/10.3109/02699201003710606>
- van de Riet, E., & Slot, P. (2017). *Structural, Process and Curriculum Quality of Day Care and Preschool Provisions in the Netherlands*. <https://doi.org/10.18778/2450-4491.05.10>
- Vitiello, V. E., Booren, L. M., Downer, J. T., & Williford, A. P. (2012). Variation in children's classroom engagement throughout a day in preschool: Relations to classroom and child factors. *Early Childhood Research Quarterly*, 27, 210–220. <https://doi.org/10.1016/j.ecresq.2011.08.005>
- Vitiello, V. E., Moas, O., Henderson, H. A., Greenfield, D. B., & Munis, P. M. (2012). Goodness of fit between children and classrooms: Effects of child temperament and

- preschool classroom quality on achievement trajectories. *Early Education and Development*, 23, 302–322. <https://doi.org/10.1080/10409289.2011.526415>
- Von Suchodoletz, A., Gunzenhauser, C., & Larsen, R. A. (2015). Die Beobachtung von Interaktionen im Kindergartenalltag. *Frühe Bildung*, 4, 211–217. <https://doi.org/10.1026/2191-9186/a000207>
- Williams, S. T., Mastergeorge, A. M., & Ontai, L. L. (2010). Caregiver involvement in infant peer interactions: Scaffolding in a social context. *Early Childhood Research Quarterly*, 25, 251–266. <https://doi.org/10.1016/j.ecresq.2009.11.004>
- Williford, A. P., Vick Whittaker, J. E., Vitiello, V. E., & Downer, J. T. (2013). Children's engagement within the preschool classroom and their development of self-regulation. *Early Education and Development*, 24, 162–187. <https://doi.org/10.1080/10409289.2011.628270>
- Wyslowska, O., & Slot, P. L. (2020). Structural and process quality in early childhood education and care provisions in Poland and the Netherlands: A cross-national study using cluster analysis. *Early Education and Development*, 31, 524–540. <https://doi.org/10.1080/10409289.2020.1734908>